Supplemental Material

The ISW1 and CHD1 ATP-dependent chromatin remodelers compete to set nucleosome spacing *in vivo*

Josefina Ocampo[*], Răzvan V. Chereji[*], Peter Eriksson, and David J. Clark[†]

Program in Genomics of Differentiation, *Eunice Kennedy Shriver* National Institute for Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892, USA

[*]These authors contributed equally to this work.
[†]Corresponding author. E-mail: clarkda@mail.nih.gov

Contents

Experimental Procedures

Detection of NDRs and flanking nucleosomes

In order to detect the position of the nucleosome depleted region (NDR) and of the two flanking nucleosomes (+1 and -1) for every gene, we use the following algorithm. We start with the raw nucleosome dyads profile, which we smooth with a moving average filter with a span of 101 bp. Let us denote this smoothed profile by y. For every gene we select the window [TSS - 1000, TSS + 1000], where we detect the peaks separated by at least 100 bp, using the `findpeaks` MATLAB function, with a minimum threshold of 0.6 times the average of y in this window. After we detected the nucleosomal peaks, we compute the width of the gaps between neighboring peaks. We identify as possible NDRs, the extended gaps, that have a width of at least 190 bp and the center between TSS - 1000 bp and TSS + 200 bp. For every extended gap, we assign a score,

$$S_i = \frac{y(\text{gap}_i)}{\frac{y(\text{peak}_i)+y(\text{peak}_{i+1})}{2}},$$

where $y(\text{peak}_i)$ represents the smoothed dyad density at the location of peak i, and $y(\text{gap}_i)$ represents the smoothed dyad density in the middle of gap i, between peaks i and $i+1$. This score represents the fraction of nucleosome density present in the middle of two nucleosomes, relative to the average density at the nucleosomal peak centers. If two extended gaps are found close to TSS (*i.e.* the middle of the gap is situated less than 200 bp away from TSS), then we pick the gap with the lowest score as the real NDR corresponding to that gene; otherwise, we pick the extended gap which is closest to TSS. After the NDR is detected, the nearest peaks on both sides of it are called the +1 and -1 nucleosomes (Suplemental Figure S3A), depending on the gene direction. We detect all NDRs and all pairs of flanking nucleosomes using data from two replicates. If the coordinates determined in the two replicates differed by more than 50 bp, then the corresponding genes are manually inspected using the IGV browser (Robinson et al., 2011), and these coordinates are manually curated.

Estimation of nucleosome spacing

In order to estimate the spacing in the nucleosome arrays that are formed downstream of TSSs, the following algorithm is used. We start with the nucleosome dyad profile, smoothed with a moving average filter with a span of 101 bp, y. For every gene, the window $[x_{+1}-200, x_{+1}+600]$ is selected, where x_{+1} represents the coordinate of the +1 nucleosome dyad. This window contains the NDR and the first four or five nucleosomes on the gene body. We construct a periodic pattern composed of six Gaussian distributions,

$$P_d(x) = G_0(x) + G_d(x) + G_{2d}(x) + G_{3d}(x) + G_{4d}(x) + G_{5d}(x),$$

where $G_D(x)$ represents a Gaussian distribution with mean D and standard deviation of 40 bp,

$$G_D(x) = e^{-\frac{(x-D)^2}{2\sigma^2}},$$
$$\sigma = 40.$$

We compute the cross-correlation between the smoothed dyad density $y([x_{+1}-200, x_{+1}+600])$ and the periodic pattern P_d([-200, 600]), with a lag s, $C(d,s)$, and we select the pair $\{d,s\}$ which

maximizes the cross-correlation. Using this method, we obtain an estimation for the nucleosome spacing, d, and the shift of the nucleosome array, s, for every gene (Suplemental Figure S2B).

Estimation of nucleosome phasing

We proposed before (Ganguli et al., 2014) a simple model for the action of spacing enzymes, that can explain the formation of phased nucleosomes at the 5' ends of the genes. If the spacing enzymes start from a barrier located in the NDR, and travel along the DNA and regularly space nucleosomes (with an uncertainty characterized by the standard deviation of a Gaussian distribution), then the predicted distribution of nucleosome dyads agrees very well with the real distribution observed in experiments (Fig. S2A). The standard deviation of the fitted Gaussian reflects the effective phasing potential of all spacing enzymes that are acting in the given sample. If the spacing enzymes act consistently while creating the regular nucleosome arrays (with a very small variability), then the arrays created in different cells and on different genes will be in phase. On the other hand, if the spacing enzyme has a big variability in creating the nucleosome arrays, then different genes and different cells will contain variable configurations of nucleosomes, and when we compute the genome-wide average distribution of nucleosomes we observe poor phasing (small and broad peaks corresponding to nucleosomes +1, +2, etc.).

We fitted the nucleosome distributions at the 5' ends of the yeast genes, for wild type cells and different remodeler mutants (Fig. S2A). We observe that wild type cells present the best phased nucleosome arrays (σ = 14.4 and 16.0 in the 2 replicates). The next best phased arrays are observed in *isw2*Δ cells (σ = 15.3 and 16.1). Among the single mutants, the worst phasing is observed in *isw1*Δ cells (σ = 18.4 and 18.9) and *chd1*Δ cells (σ = 19.5 and 20.2). The double mutant *isw1*Δ *chd1*Δ has even worse phasing (σ = 27.5 and 28.1).

We also checked how the nucleosome phasing varies among genes with different nucleosome spacing. We separated all yeast genes into 5 quantiles, from the shortest spacing (quantile 1) to the longest spacing (quantile 5), and we compared the overall phasing in these groups of genes (Fig. S2B). We see that in general the genes with extreme spacing have the weakest phasing, which is expected since these contain the genes that are most transcribed.

Paired-end sequencing of chromatin immunoprecipitates (PESCI)

The method was as described (Ganguli et al., 2014; Cole et al., 2014), with some modifications: Cells were fixed with 1% formaldehyde (15 min at room temperature) and glycine was added to 0.4 M for 5 min. Cells were collected by filter, washed with 20 mM Tris-HCl pH 8.0, 0.15 M NaCl, aliquoted in batches of ~70 OD_{600} units and stored at -80°C. A frozen cell pellet was resuspended in 0.4 ml FA Buffer (50 mM HEPES-KOH pH 7.5, 0.14 M NaCl, 1 mM EDTA, 1% Triton-X100, 0.1% sodium deoxycholate) with protease inhibitors (Roche 05 056 489 001). Cells were vortexed in the presence of an equal volume of glass beads (0.5 mm soda-lime BioSpec 11079105) for 40 min at 4°C. The beads were removed. The samples were sonicated using a Diagenode Bioruptor 300 at 4°C with 40 cycles of 30 s high power and 30 s rest, with a 15 min cooling period after 20 cycles. Debris was removed by spinning for 20 min at 14,000 rpm and 4°C. The supernatant was aliquoted and stored at -80°C. DNA was quantified using the Hoechst assay. Anti-Rpb3 (1.2 μg Neoclone WP012) was mixed with 50 μl beads (1:1 slurry) (Invitrogen Dynabeads Pan Mouse IgG 110.42) that had been washed twice with 0.5 ml 5 mg/ml BSA in 1x PBS and incubated for 3 h at 4°C. Controls without antibody were also done. For IP, 10 μg DNA was adjusted to 400 μl with

FA buffer containing 0.75 mg BSA/ml. Binding was performed with rotation for 3 h at 4°C. The supernatant was removed. The beads were washed twice with FA buffer containing 0.025% SDS, once with FA buffer, twice with 50 mM HEPES-KOH pH 7.5, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, twice with 10 mM Tris-HCl pH 8.0, 250 mM LiCl, 1 mM EDTA, 0.5% NP40, 0.5% sodium deoxycholate and once with TE. Bound DNA was eluted with 100 μl 50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% SDS, 0.14 M NaCl in a Thermomixer (15 min, 1400 rpm, 65°C). The eluate was removed, the elution was repeated with 150 μl 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.67% SDS, 0.14 M NaCl and the eluates were pooled. Cross-links were reversed by incubation at 65°C overnight, followed by proteinase K digestion (0.4 mg/ml, 2 h at 37°C and 2 h at 55°C). DNA extraction and library preparation were performed as above. Input DNA was prepared as described (Cole et al., 2014).

Software and external data used in this study

Paired-end reads were aligned using Bowtie 2 (Langmead and Salzberg, 2012) against the UCSC sacCer3 genome assembly. Raw coverage profiles were generated using BEDTools utilities (Quinlan and Hall, 2010), and viewed in IGV (Robinson et al., 2011). For further analysis, nucleosome sequences in the range 120–160 bp were selected, and the locations of their dyads were inferred by calculating the midpoint coordinate. The sequencing depths for all data sets were adjusted to the common value of 1 read per bp. Transcript end coordinates (TSS and TTS) were obtained from Park et al. (2014). All data analysis was performed in MATLAB.

Yeast strains

Mating, sporulation and tetrad analysis were performed as described in Sherman and Hicks (1991). YTT186 and YTT196 (Tsukiyama et al., 1999) were crossed to generate a diploid strain, which was transformed with a PCR fragment, resulting in the insertion of a GAL1 promoter and an N-terminal 3x HA tag at the *RSC8* locus. The PCR fragment was made using pFA6a-kanMX6-pGAL1-3HA (Longtine et al., 1998) and genomic DNA as templates in a three-step PCR with primers 1756, 1757, 1758, 1759 and 1764, including an upstream extension of RSC8 to improve integration efficiency.

Forward primer 1756 (*RSC8* promoter / kanamycin resistance gene): AGAGAGAAGGGTAGAGCAGAGAAGCAAAGAGTAGATAGAAGGCCAGGATAGAATTCGAGCTCGTTTAAAC
Reverse primer 1757 (*RSC8* coding region): TCGTGCGAGTCTACCATAGGGACATCCTTATCCTTTTCAGTGTCGCTCATGCACTGAGCAGCGTAATCTG
Reverse primer 1758 (*RSC8* promoter): TATCCTGGCCTTCTATCTACTCTTTG
Forward primer 1759 (1.3 kb upstream of *RSC8*): ATTTTGTAAAGGCGTCCCTGT
Reverse primer 1764 (*RSC8* coding region): TTCCTCCTTCGCCTGTTCC

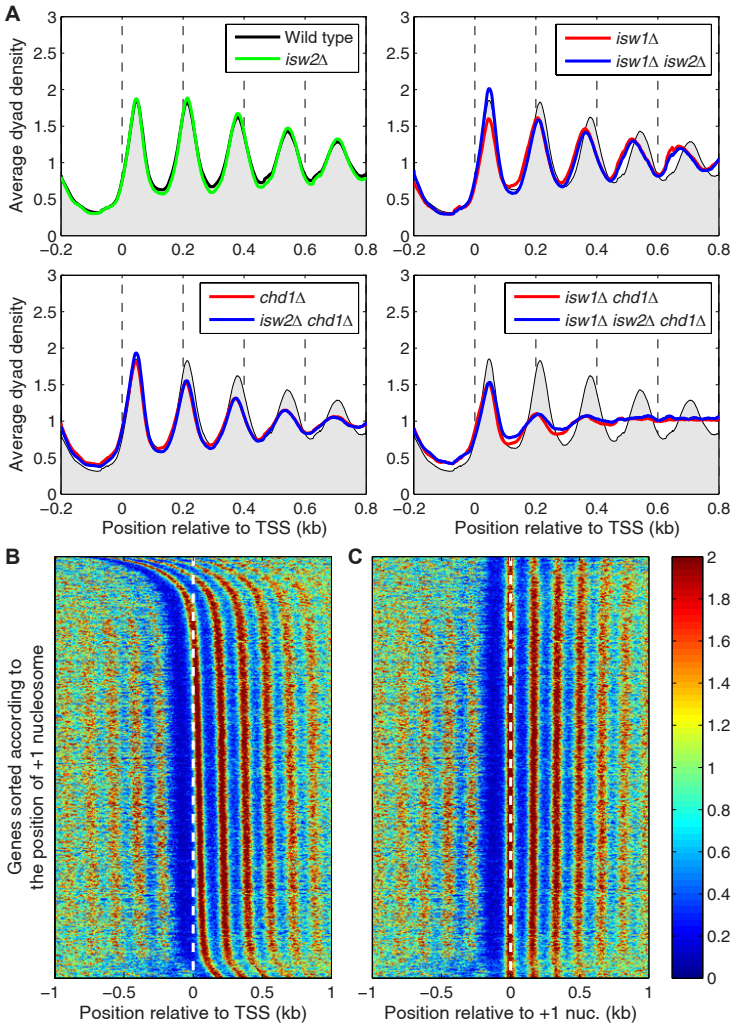
GAL-3HA-RSC8 transformants were selected on galactose plates with 200 μg/ml G418. This transformation resulted in the diploid strain YJO504. Dissection of YJO504 resulted in haploid strains YPE458, YPE460 and YPE470. Diploid strain YJO504 was transformed with a PCR fragment to delete *CHD1*. This PCR fragment was generated in a three-step PCR using pFA6a-His3Mx6 (Longtine et al., 1998) and genomic DNA as templates with primers 1796-1799; the integration fragment was extended upstream to improve integration.

Forward primer 1796 (*CHD1* promoter/ *HIS5*) TTTCTTTAAACCTATACCCAATTCAAAGCAGAACCTTTTCTAATTTAATTCTCACTTATACGGATCCCCGGGTTAATTAA
Reverse primer 1797 (*CHD1* coding region): AAATACGTTTATAGTTATGGGGGGAAGGAA-

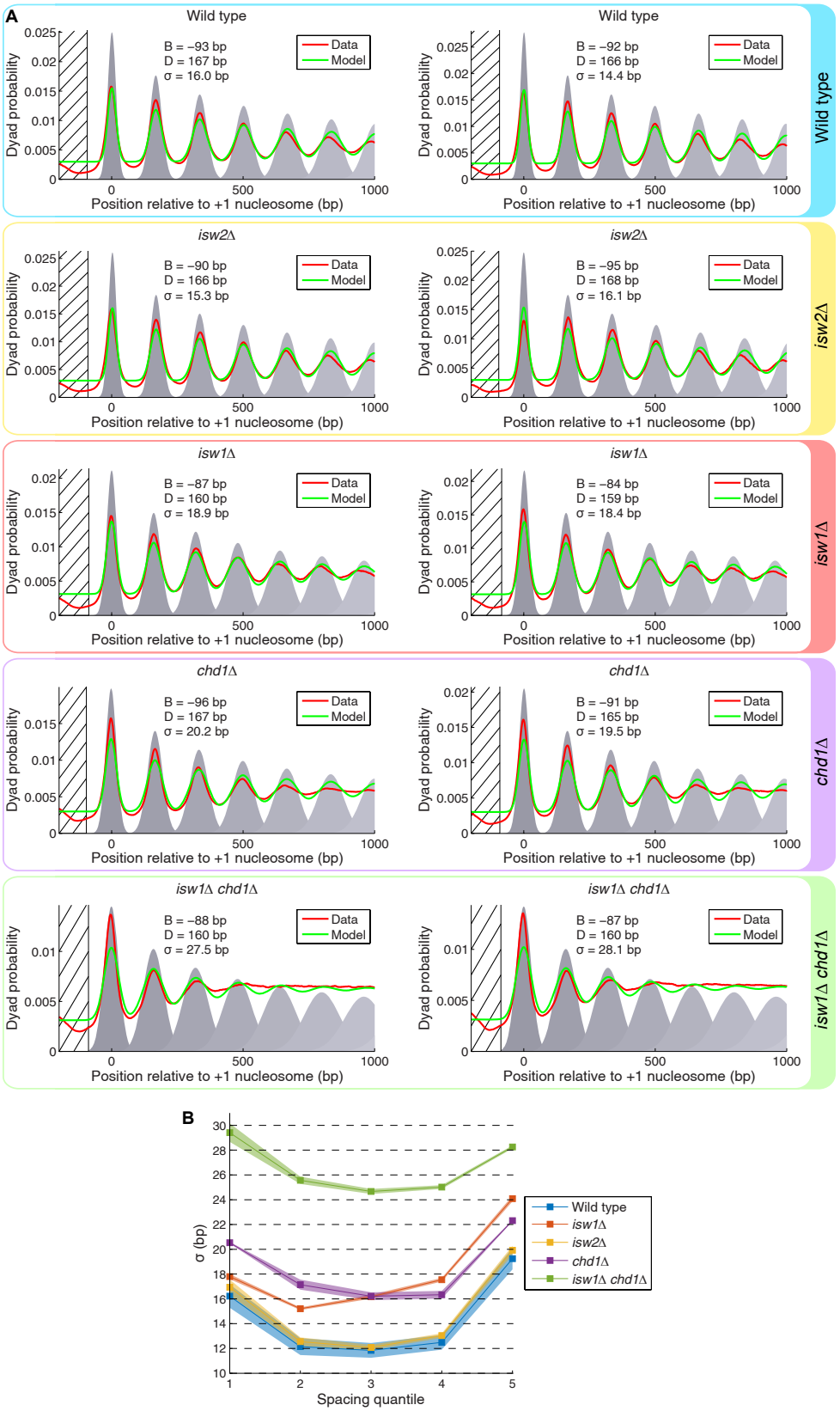
CAATGGAAAATGTGGTGAAGAAAAATTGTTGAATTCGAGCTCGTTTAAAC
Forward primer 1798 (890 bp upstream of *CHD1*): CCCATATACGGTAAAAGATCATACGTTG
Reverse primer 1799 (*CHD1* promoter): TAAGTGAGAATTAAATTAGAAAAGGTTCTGCTT

This transformation resulted in diploid strain YJO505. Dissection of YJO505 gave the haploid strains YJO482, YJO484, YJO486 and YJO488. Genotypes are listed in Table S1.

Supplemental Figures

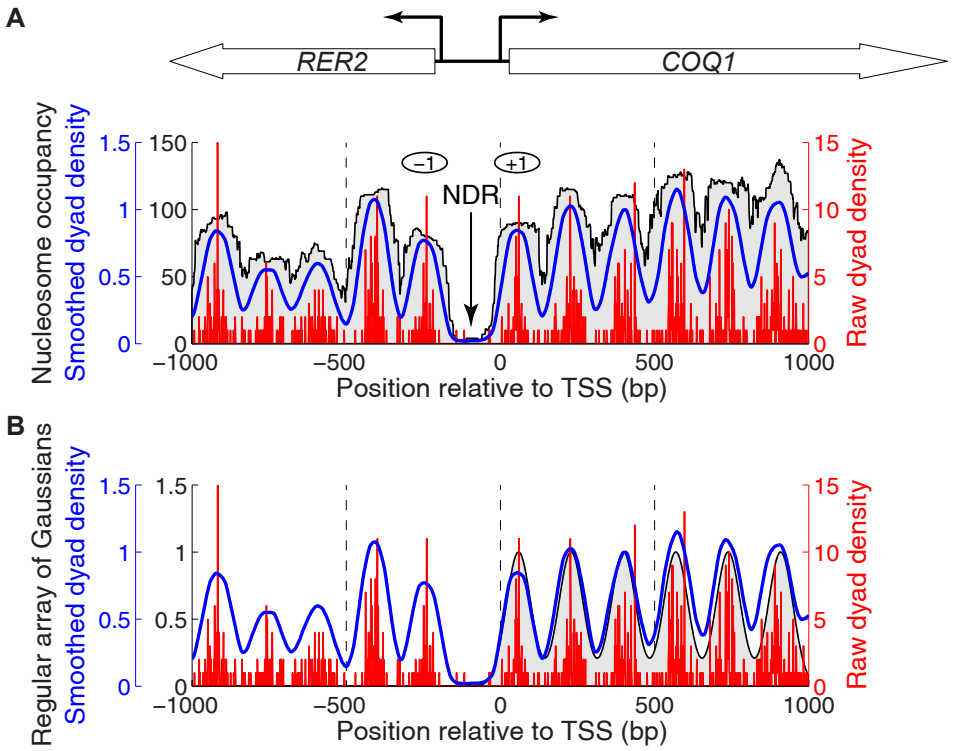


Supplemental Figure S1. Alignment of yeast genes on the +1 nucleosome for spacing analysis. (A) Nucleosome phasing profiles for all genes aligned on the transcription start site (TSS) in wild type cells. Phasing plots for isogenic strains corresponding to all combinations of the *chd1*Δ, *isw1*Δ and *isw2*Δ mutations. The sequencing depths for all data sets were adjusted to 1 read per bp. Wild type is shown as a grey background in all panels. Re-alignment on the dyad of the +1 nucleosome in wild type cells results in sharper peaks with larger amplitudes (compare with Figure 1B). Heat map analysis of all genes showing the nucleosome dyad distributions in 2-kb windows centered on (B) the TSS (white dashed line), or on (C) the +1 nucleosome (white dashed line). Each row represents a gene. In both heat maps, the genes are sorted according to the position of the +1 nucleosome relative to the TSS, which increases from top to bottom. The heat maps were smoothed with a 2D symmetrical Gaussian filter with standard deviation $\sigma = 3$. When all genes are sorted according to the location of the NDR relative to the TSS in wild type cells, it is apparent that the NDR is not always located just upstream of the TSS. Most genes have an NDR immediately upstream of the TSS, but for some genes, the NDR is located far upstream, or even downstream of the TSS. To eliminate this problem, we re-aligned the nucleosome arrays relative to the +1 nucleosome instead of the TSS. This re-alignment allows us to obtain a meaningful average spacing, which is not affected by interference from genes that have displaced nucleosome arrays, such as those at the top in B.

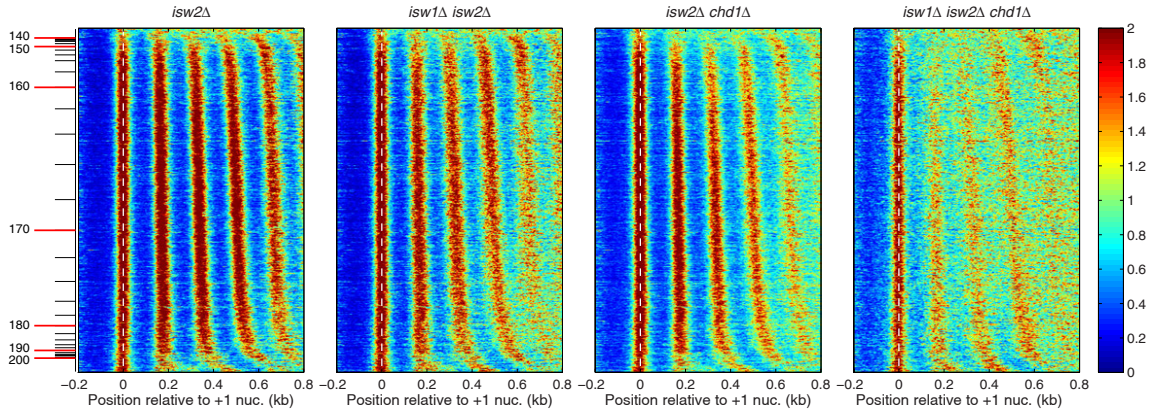


Supplemental Figure S2. Caption on next page.

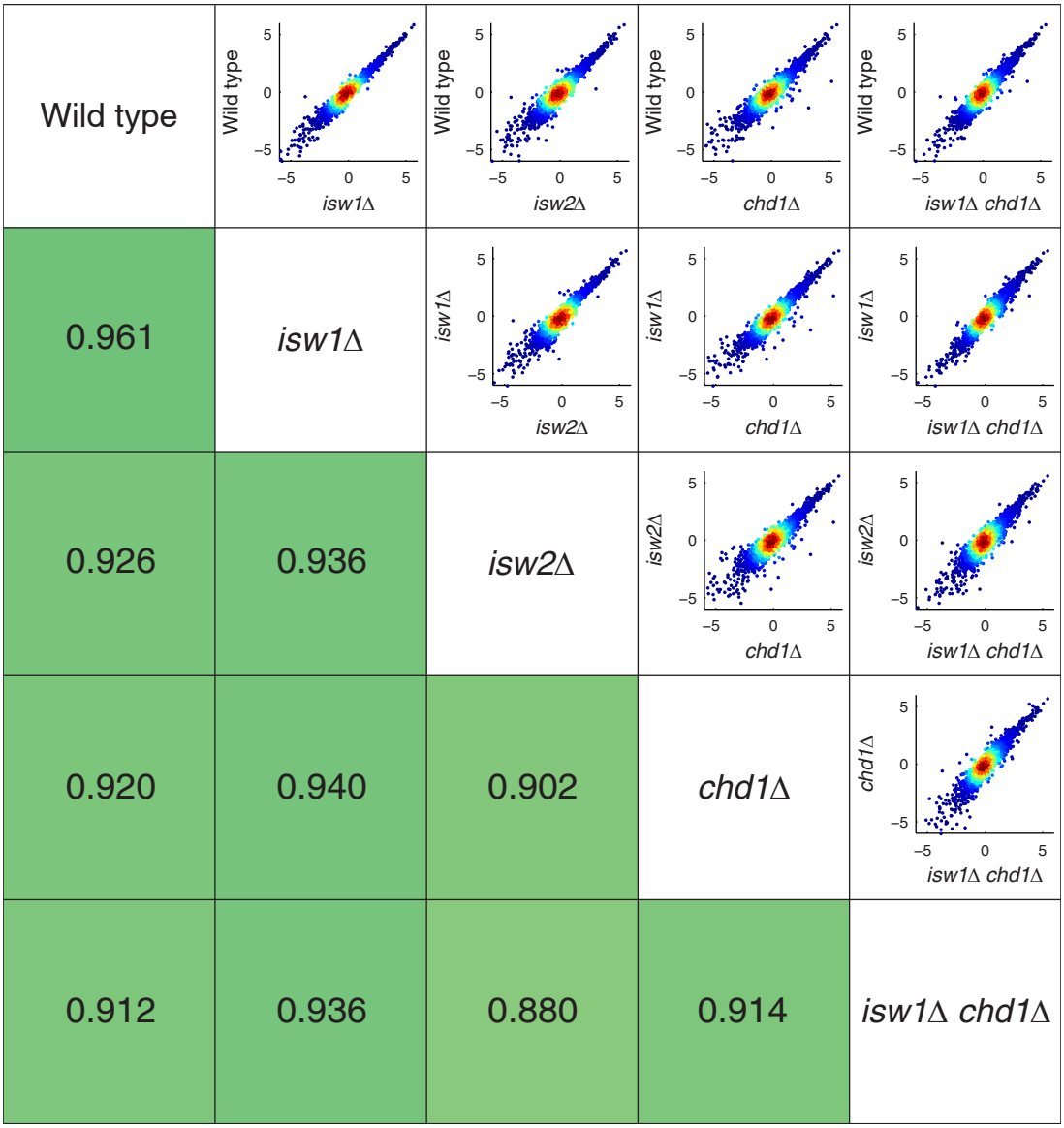
Supplemental Figure S2. Estimation of nucleosome phasing in remodeler mutants using a simple barrier model. Using the model proposed before (Ganguli et al., 2014), we quantify the degree of nucleosome phasing (see Experimental Procedures above). (A) Overall phasing in wild type cells and in *isw2*Δ, *isw1*Δ, *chd1*Δ and *isw1*Δ *chd1*Δ strains. We also separated the yeast genes into 5 quantiles, from the shortest spacing (quantile 1) to the longest spacing (quantile 5), and we compared the overall phasing in these groups of genes. (B) Comparison of the phasing among groups of genes with different spacing.



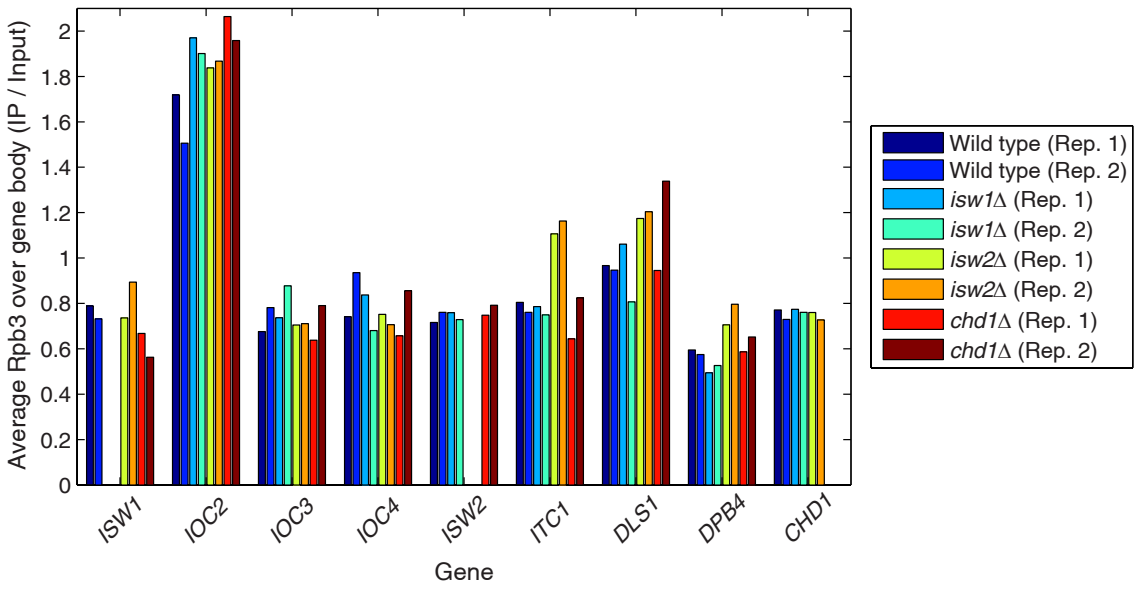
Supplemental Figure S3. Detection of +1/-1 nucleosomes and estimation of nucleosome spacing. (A) Example of a typical nucleosome organization, near the promoter of COQ1 gene: an NDR close to TSS, ranked by regular arrays of nucleosomes on both sides. Raw nucleosome dyad counts (red line), smoothed dyad density (blue line) and nucleosome occupancy profile (grey background area) are shown. Using the smoothed dyad profile, we detect the locations of the NDR (black arrow), and of the pair of bordering nucleosomes: +1 and -1 (black ellipses). (B) By maximizing the cross-correlation between the smoothed dyads profiles (blue line) and a set of periodic Gaussian distributions (grey background area), we estimate the periodicity and location of the regular arrays of nucleosomes from the gene bodies.



Supplemental Figure S4. ISW2 has little effect on nucleosome spacing at the global level. Heat map analysis of all genes showing nucleosome dyad distributions, aligned on the +1 nucleosome dyad. Each row represents a gene. Genes are sorted in order of increasing average spacing in wild type cells, from top to bottom. The color scale represents nucleosomal dyad density (red: high; blue: low dyad density). Equivalent heat maps for *ISW2* strains are shown in Figure 2A.



Supplemental Figure S5. Gene expression patterns in remodeling mutants, measured by Pol II occupancy. The average Rpb3 occupancy obtained from paired-end ChIP-seq experiments was calculated for the coding region of each gene in wild type, all three single mutants and the *isw1*Δ *chd1*Δ double mutant. The average value for each gene obtained from two biological replicate experiments was used in the scatter plots. Each gene is represented by a blue dot, but where many dots overlap, color is used to indicate the spot density (red: high density; yellow: lower density). Below the main diagonal, the Spearman's rank correlation coefficients are provided for comparison of wild type and mutants with one another.



Supplemental Figure S6. Average Rpb3 occupancy on genes encoding the subunits of the three remodeling complexes: ISW1 (*ISW1, IOC2, IOC3, IOC4*), ISW2 (*ISW2, ITC1, DLS1, DPB4*) and CHD1 (*CHD1*). When one of these genes was deleted (*isw1*Δ - light blue and turquoise bars, *isw2*Δ - yellow and orange bars, *chd1*Δ - red and maroon bars), the expression of the genes encoding the other remodelers remained virtually unaffected.

Supplemental Tables

Supplemental Table S1. Genotypes and sources of yeast strains used in this study.

Strain	Genotype	Source
YTT186	*MATa ade2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 RAD5+ isw1Δ::ADE2*	Tsukiyama et al. (1999)
YTT196	*MATα ade2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 RAD5+ isw2Δ::LEU2*	Tsukiyama et al. (1999)
YJO504	*MATa/MATα ade2-1/ade2-1 can1-100/can1-100 his3-11,15/his3-11,15 leu2-3,112/leu2-3,112 trp1-1/trp1-1 ura3-1/ura3-1 RAD5+/RAD5+ ISW1/isw1Δ::ADE2 ISW2/isw2Δ::LEU2 rsc8::KanMX-GALp-HA3-RSC8*	This study
YJO505	*MATa/MATα ade2-1/ade2-1 can1-100/can1-100 his3-11,15/his3-11,15 leu2-3,112/leu2-3,112 trp1-1/trp1-1 ura3-1/ura3-1 RAD5+/RAD5+ ISW1/isw1Δ::ADE2 ISW2/isw2Δ::LEU2 RSC8/rsc8::KanMX-GALp-HA3-RSC8 CHD1/chd1Δ::HIS5*	This study
YPE458	*MATa ade2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 RAD5+*	This study
YPE460	*MATa ade2-1 can1-100 his3-11,15 leu2-3,112 trp1-1ura3-1 RAD5+ isw1Δ::ADE2 isw2Δ::LEU2*	This study
YPE470	*MATa ade2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 RAD5+ isw2Δ::LEU2*	This study
YJO482	*MATa ade2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 RAD5+ chd1Δ::HIS5*	This study
YJO484	*MATa ade2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 RAD5+ isw1Δ::ADE2 chd1Δ::HIS5*	This study
YJO486	*MATa ade2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 RAD5+ isw2Δ::LEU2 chd1Δ::HIS5*	This study
YJO488	*MATa ade2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 RAD5+ isw1Δ::ADE2 isw2Δ::LEU2 chd1Δ::HIS5*	This study

Supplemental Table S2. Global average nucleosome spacing in the remodeling mutants. Averages and standard deviations derive from two biological replicate experiments. The spacing was determined by regression analysis of peak values for the first five nucleosomes on all genes after alignment on the +1 nucleosome. *ND* – not determined because phasing is too poor.

Sample	Average spacing	Standard deviation
Wild type	166	0
*isw1*Δ	159	1
*isw2*Δ	166	0
*chd1*Δ	164	1
*isw1*Δ *isw2*Δ	159	1
*isw1*Δ *chd1*Δ	*ND*	*ND*
*isw2*Δ *chd1*Δ	164	1
*isw1*Δ *isw2*Δ *chd1*Δ	*ND*	*ND*

Supplemental Table S3. Summary of paired-end data – Nucleosome mapping (MNase-seq) and Rpb3 ChIP-seq (PESCI). All reads were 50 nt long.

MNase-seq sample	Yeast strain	No. of aligned reads
Wild type (replicate 1)	YPE458	73,496,483
Wild type (replicate 2)	YPE458	8,084,621
*isw1*Δ (replicate 1)	YTT186	7,872,706
*isw1*Δ (replicate 2)	YTT186	19,840,925
*isw2*Δ (replicate 1)	YPE470	22,245,814
*isw2*Δ (replicate 2)	YPE470	24,681,479
*chd1*Δ (replicate 1)	YJO482	21,592,761
*chd1*Δ (replicate 2)	YJO482	21,163,843
*isw1*Δ *isw2*Δ (replicate 1)	YPE460	21,759,297
*isw1*Δ *isw2*Δ (replicate 2)	YPE460	7,407,759
*isw1*Δ *chd1*Δ (replicate 1)	YJO484	11,828,476
*isw1*Δ *chd1*Δ (replicate 2)	YJO484	25,532,960
*isw2*Δ *chd1*Δ (replicate 1)	YJO486	22,128,418
*isw2*Δ *chd1*Δ (replicate 2)	YJO486	16,820,906
*isw1*Δ *isw2*Δ *chd1*Δ (replicate 1)	YJO488	13,183,289
*isw1*Δ *isw2*Δ *chd1*Δ (replicate 2)	YJO488	24,294,504

Rpb3 ChIP-seq sample	Yeast strain	No. of aligned reads
Wild type Input (replicate 1)	YPE458	23,459,761
Wild type IP (replicate 1)	YPE458	12,446,188
Wild type Input (replicate 2)	YPE458	6,712,213
Wild type IP (replicate 2)	YPE458	12,829,278
*isw1*Δ Input (replicate 1)	YTT186	21,623,664
*isw1*Δ IP (replicate 1)	YTT186	27,369,474
*isw1*Δ Input (replicate 2)	YTT186	10,302,008
*isw1*Δ IP (replicate 2)	YTT186	8,586,733
*isw2*Δ Input (replicate 1)	YPE470	10,607,785
*isw2*Δ IP (replicate 1)	YPE470	8,459,811
*isw2*Δ Input (replicate 2)	YPE470	2,271,556
*isw2*Δ IP (replicate 2)	YPE470	7,036,973
*chd1*Δ Input (replicate 1)	YJO482	16,999,250
*chd1*Δ IP (replicate 1)	YJO482	8,080,351
*chd1*Δ Input (replicate 2)	YJO482	11,760,467
*chd1*Δ IP (replicate 2)	YJO482	4,037,297
*isw1*Δ *chd1*Δ Input (replicate 1)	YJO484	17,456,734
*isw1*Δ *chd1*Δ IP (replicate 1)	YJO484	18,916,825
*isw1*Δ *chd1*Δ Input (replicate 2)	YJO484	8,616,192
*isw1*Δ *chd1*Δ IP (replicate 2)	YJO484	20,540,227

Supplemental References

Cole HA, Ocampo J, Iben JR, Chereji RV, and Clark DJ. 2014. Heavy transcription of yeast genes correlates with differential loss of histone H2B relative to H4 and queued RNA polymerases. *Nucleic Acids Res.* **42**: 12512–12522.

Ganguli D, Chereji RV, Iben JR, Cole HA, and Clark DJ. 2014. Rsc-dependent constructive and destructive interference between opposing arrays of phased nucleosomes in yeast. *Genome research* **24**: 1637–1649.

Langmead B and Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**: 357–359.

Longtine MS, Mckenzie III A, Demarini DJ, Shah NG, Wach A, Brachat A, Philippsen P, and Pringle JR. 1998. Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast* **14**: 953–961.

Park D, Morris AR, Battenhouse A, and Iyer VR. 2014. Simultaneous mapping of transcript ends at single-nucleotide resolution and identification of widespread promoter-associated non-coding RNA governed by TATA elements. *Nucleic Acids Res.* **42**: 3736–3749.

Quinlan AR and Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–842.

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, and Mesirov JP. 2011. Integrative genomics viewer. *Nat. Biotechnol.* **29**: 24–26.

Sherman F and Hicks J. 1991. Micromanipulation and dissection of asci. *Methods in Enzymology* **194**: 21–37.

Tsukiyama T, Palmer J, Landel CC, Shiloach J, and Wu C. 1999. Characterization of the imitation switch subfamily of ATP-dependent chromatin-remodeling factors in Saccharomyces cerevisiae. *Genes Dev.* **13**: 686–697.